

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: ANDERSON, Darrell R.  
HANNA, Nabil  
LEONARD, John E.  
NEWMAN, Roland A.  
REFF, Mitchell E.  
RASTETTER, William H.

(ii) TITLE OF INVENTION: THERAPEUTIC APPLICATION OF CHIMERIC AND  
RADIOLABELED ANTIBODIES TO HUMAN B LYMPHOCYTE RESTRICTED  
DIFFERENTIATION ANTIGEN FOR THE TREATMENT OF B CELL  
LYMPHOMA

(iii) NUMBER OF SEQUENCES: 9

(iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
(B) STREET: P.O. Box 1404  
(C) CITY: Alexandria  
(D) STATE: Virginia  
(E) COUNTRY: United States  
(F) ZIP: 22313-1404

(v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: US 08/149,099  
(B) FILING DATE: 03-NOV-1993  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 07/978,891  
(B) FILING DATE: 12-NOV-1992

(viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Teskin, Robin L.  
(B) REGISTRATION NUMBER: 35,030  
(C) REFERENCE/DOCKET NUMBER: 012712-014

(ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: (703) 836-6620  
(B) TELEFAX: (703) 836-2021

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGGAGCTTGG ATCGATCCTG TATGGTT

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 8540 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GACGTCGCGG CCGCTCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG	60
AGGCCGAGGC GGCCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGCATGGGGC	120
GGAGAATGGG CGGAACCTGGG CGGAGTTAGG GGCAGGATGG GCGGAGTTAG GGGCGGGACT	180
ATGGTTGCTG ACTAATTGAG ATGCATGCTT TGCACTACTTC TGCCTGCTGG GGAGCCTGGG	240
GACTTTCCAC ACCTGGTTGC TGACTAATTG AGATGCATGC TTTGCATACT TCTGCCTGCT	300
GGGGAGCCTG GGGACTTTCC ACACCCCTAAC TGACACACAT TCCACAGAAAT TAATTCCCCT	360
AGTTATTAAT AGTAATCAAT TACGGGTCA TTAGTTCAT A GCCCATATAT GGAGTTCCGC	420
GTTACATAAC TTACGGTAAA TGGCCCGCT GGCTGACCGC CCAACGACCC CCGCCCATTG	480
ACGTCAATAA TGACGTATGT TCCCAGTAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA	540
TGGGTGGACT ATTTACGGTA AACTGCCAC TTGGCAGTAC ATCAAGTGTAA TCATATGCCA	600
AGTACGCCCT CTATTGACGT CAATGACGGT AAATGGCCCG CCTGGCATTAA TGCCCAGTAC	660
ATGACCTTAT GGGACTTTCC TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC	720
ATGGTGATGC GGTTTTGGCA GTACATCAAT GGGCGTGGAT AGCGGTTGA CTCACGGGA	780
TTTCCAAGTC TCCACCCCCAT TGACGTCAAT GGGAGTTGT TTTGGCACCA AAATCAACGG	840
GACTTTCCAA AATGTGTTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGGGTGTA	900
CGGTGGGAGG TCTATATAAG CAGAGCTGGG TACGTGAACC GTCAGATCGC CTGGAGACGC	960
CATCACAGAT CTCTCACCAT GAGGGTCCCC GCTCAGCTCC TGGGCTCCT GCTGCTCTGG	1020
CTCCCAGGTG CACGATGTGA TGGTACCAAG GTGGAAATCA AACGTACGGT GGCTGCACCA	1080
TCTGTCTTCA TCTTCCCGCC ATCTGATGAG CAGTTGAAAT CTGGAACTGC CTCTGTTGTG	1140
TGCCTGCTGA ATAACCTCTA TCCCAGAGAG GCCAAAGTAC AGTGGAAAGGT GGATAACGCC	1200
CTCCAATCGG GTAACCTCCA GGAGAGTGTG ACAGAGCAGG ACAGCAAGGA CAGCACCTAC	1260
AGCCTCAGCA GCACCCCTGAC GCTGAGCAAA GCAGACTACG AGAAACACAA AGTCTACGCC	1320
TGCGAAGTCA CCCATCAGGG CCTGAGCTCG CCCGTACAA AGAGCTTCAA CAGGGGAGAG	1380
TGTTGAATTG AGATCCGTTA ACGGTTACCA ACTACCTAGA CTGGATTCTG GACAACATGC	1440
GGCCGTGATA TCTACGTATG ATCAGCCTCG ACTGTGCCTT CTAGTTGCCA GCCATCTGTT	1500
GGTTGCCCCCT CCCCCGTGCC TTCCCTTGACC CTGGAAGGTG CCACCTCCAC TGTCCCTTCC	1560

AGACAGTGT	CTCTGCACAG	ATAAGGACAA	ACATTATTCA	GAGGGAGTAC	CCAGAGCTGA	3660
GACTCCTAAG	CCAGTGAGTG	GCACAGCATT	CTAGGGAGAA	ATATGCTTGT	CATCACCGAA	3720
GCCTGATTCC	GTAGAGCCAC	ACCTTGGTAA	GGGCCAATCT	GCTCACACAG	GATAGAGAGG	3780
GCAGGAGCCA	GGGCAGAGCA	TATAAGGTGA	GGTAGGATCA	GTTGCTCTC	ACATTTGCTT	3840
CTGACATAGT	TGTGTTGGGA	GCTTGGATAG	CTTGGACAGC	TCAGGGCTGC	GATTCGCGC	3900
CAAACTTGAC	GGCAATCCTA	CGGTGAAGGC	TGGTAGGATT	TTATCCCCGC	TGCCATCATG	3960
GTTCGACCAT	TGAACTGCAT	CGTCGCCGTG	TCCAAAATA	TGGGGATTGG	CAAGAACGGA	4020
GACCTACCCCT	GGCCTCCGCT	CAGGAACGAG	TTCAAGTACT	TCCAAAGAAT	GACCACAACC	4080
TCTTCAGTGG	AAGGTAAACA	GAATCTGGTG	ATTATGGTA	GGAAAACCTG	GTTCTCCATT	4140
CCTGAGAACCA	ATCGACCTTT	AAAGGACAGA	ATTAATATAG	TTCTCAGTAG	AGAACTCAAA	4200
GAACCACAC	GAGGAGCTCA	TTTTCTTGCC	AAAAGTTGG	ATGATGCCTT	AAGACTTTATT	4260
GAACAACCGG	AATTGGCAAG	TAAAGTAGAC	ATGGTTGGA	TAGTCGGAGG	CAGTTCTGTT	4320
TACCAGGAAG	CCATGAATCA	ACCAGGCCAC	CTTAGACTCT	TTGTGACAAG	GATCATGCAG	4380
GAATTGAAA	GTGACACGTT	TTTCCCAGAA	ATTGATTTGG	GGAAATATAA	ACTTCTCCCA	4440
GAATACCCAG	GCGCCTCTC	TGAGGTCCAG	GAGGAAAAAG	GCATCAAGTA	TAAGTTGAA	4500
GTCTACGAGA	AGAAAGACTA	ACAGGAAGAT	GCTTCAAGT	TCTCTGCTCC	CCTCCTAAAG	4560
TCATGCATTT	TTATAAGACC	ATGGGACTTT	TGCTGGCTTT	AGATCAGCCT	CGACTGTGCC	4620
TTCTAGTTGC	CAGCCATCTG	TTGTTGCC	CTCCCCGTG	CCTTCCTTGA	CCCTGGAAAGG	4680
TGCCACTCCC	ACTGTCCTTT	CCTAATAAA	TGAGGAAATT	GCATCGCATT	GTCTGAGTAG	4740
GTGTCATTCT	ATTCTGGGGG	GTGGGGTGGG	GCAGGACAGC	AAGGGGGAGG	ATTGGGAAGA	4800
CAATAGCAGG	CATGCTGGGG	ATGCGTGGG	CTCTATGGAA	CCAGCTGGGG	CTCGAGCTAC	4860
TAGCTTGCT	TCTCAATTTC	TTATTGCA	AATGAGAAA	AAAGGAAAAT	TAATTTAAC	4920
ACCAATTCA	AGCAATTGCG	TTGCCAAAAA	GGATGCTTTA	GAGACAGTGT	4980	
TCTCTGCACA	GATAAGGACA	AACATTATTC	AGAGGGAGTA	CCCAGAGCTG	AGACTCCTAA	5040
GCCAGTGA	GGCACAGCAT	TCTAGGGAGA	AATATGCTTG	TCATCACCGA	AGCCTGATTC	5100
CGTAGAGCCA	CACCTTGGTA	AGGGCCAATC	TGCTCACACA	GGATAGAGAG	GGCAGGAGCC	5160
AGGGCAGAGC	ATATAAGGTG	AGTAGGATC	AGTGCTCCT	CACATTGCT	TCTGACATAG	5220
TTGTGTTGGG	AGCTTGGATC	GATCCTCTAT	GGTTGAACAA	GATGGATTGC	ACGCAGGTTTC	5280
TCCGGCCGCT	TGGGTGGAGA	GGCTATTCTGG	CTATGACTGG	GCACAAACAGA	CAATCGGCTG	5340
CTCTGATGCC	GCCGTGTTCC	GGCTGTCAGC	GCAGGGCGC	CCGGTTCTTT	TTGTCAAGAC	5400
CGACCTGTCC	GGTGCCTCTGA	ATGAACTGCA	GGACGAGGCA	GCGCGGCTAT	CGTGGCTGGC	5460
CACGACGGGC	GTTCTTGC	CAGCTGTGCT	CGACGTTGTC	ACTGAAGCGG	GAAGGGACTG	5520
GCTGCTATTG	GGCGAAGTGC	CGGGCAGGA	TCTCCTGTCA	TCTCACCTTG	CTCCTGCCGA	5580
GAAAGTATCC	ATCATGGCTG	ATGCAATGCG	GCGGCTGCAT	ACGCTTGATC	CGGCTACCTG	5640

TAATAAAATG	AGGAAATTGC	ATCGCATTGT	CTGAGTAGGT	GTCATTCTAT	TCTGGGGGGT	1620
GGGGTGGGGC	AGGACAGCAA	GGGGGAGGAT	TGGGAAGACA	ATAGCAGGCA	TGCTGGGGAT	1680
GCGGTGGGCT	CTATGGAACC	AGCTGGGCT	CGACAGCTAT	GCCAAGTACG	CCCCCTATTG	1740
ACGTCAATGA	CGGTAAATGG	CCCGCCTGGC	ATTATGCCA	GTACATGACC	TTATGGGACT	1800
TTCCTACTTG	GCAGTACATC	TACGTATTAG	TCATCGCTAT	TACCATGGTG	ATGCGGTTTT	1860
GGCAGTACAT	CAATGGGCGT	GGATAGCGGT	TTGACTCACG	GGGATTCCA	AGTCTCCACC	1920
CCATTGACGT	CAATGGGAGT	TTGTTTGGC	ACCAAAATCA	ACGGGACTTT	CCAAAATGTC	1980
GTAACAACTC	CGCCCCATTG	ACGCAAATGG	GCGGTAGGCG	TGTACGGTGG	GAGGTCTATA	2040
TAAGCAGAGC	TGGGTACGTC	CTCACATTCA	GTGATCAGCA	CTGAACACAG	ACCCGTCGAC	2100
ATGGGTTGGA	GCCTCATCTT	GCTCTTCCTT	GTCGCTGTTG	CTACGCGTGT	CGCTAGCACC	2160
AAGGGCCAT	CGGTCTTCCC	CCTGGCACCC	TCCTCCAAGA	GCACCTCTGG	GGGCACAGCG	2220
GCCCTGGGCT	GCCTGGTCAA	GGACTACTTC	CCCGAACCGG	TGACGGTGTG	GTGGAACTCA	2280
GGCGCCCTGA	CCAGCGGCGT	GCACACCTTC	CCGGCTGTCC	TACAGTCCTC	AGGACTCTAC	2340
TCCCTCAGCA	GCGTGGTGAC	CGTGCCTCTC	AGCAGCTTGG	GCACCCAGAC	CTACATCTGC	2400
AACGTGAATC	ACAAGCCCAG	CAACACCAAG	GTGGACAAGA	AAGCAGAGCC	CAAATCTTGT	2460
GACAAAACTC	ACACATGCC	ACCGTGCCCA	GCACCTGAAC	TCCTGGGGGG	ACCGTCAGTC	2520
TTCCTCTTCC	CCCCAAAACC	CAAGGACACC	CTCATGATCT	CCCGGACCCC	TGAGGTACA	2580
TGCGTGGTGG	TGGACGTGAG	CCACGAAGAC	CCTGAGGTCA	AGTTCAACTG	GTACGTGGAC	2640
GGCGTGGAGG	TGCATAATGC	CAAGACAAAG	CCGCAGGAGG	AGCAGTACAA	CAGCACGTAC	2700
CGTGTGGTCA	GCGTCCTCAC	CGTCCTGCAC	CAGGACTGGC	TGAATGGCAA	GGACTACAAG	2760
TGCAAGGTCT	CCAACAAAGC	CCTCCAGCC	CCCATCGAGA	AAACCATCTC	CAAAGCCAAA	2820
GGGCAGCCCC	GAGAACACCA	GGTGTACACC	CTGCCCCCAT	CCCGGGATGA	GCTGACCAGG	2880
AACCAAGGTCA	GCCTGACCTG	CCTGGTCAA	GGCTTCTATC	CCAGCGACAT	CGCCGTGGAG	2940
TGGGAGAGCA	ATGGGCAGCC	GGAGAACAC	TACAAGACCA	CGCCTCCCGT	GCTGGACTCC	3000
GACGGCTCCT	TCTTCCTCTA	CAGCAAGCTC	ACCGTGGACA	AGAGCAGGTG	GCAGCAGGGG	3060
AACGTCTTCT	CATGCTCCGT	GATGCATGAG	GCTCTGCACA	ACCAACTACAC	GCAGAAGAGC	3120
CTCTCCCTGT	CTCCGGGTAA	ATGAGGATCC	GTAAACGGTT	ACCAACTACC	TAGACTGGAT	3180
TCGTGACAAC	ATGCGGCCGT	GATATCTACG	TATGATCAGC	CTCGACTGTG	CCTCTAGTT	3240
GCCAGCCATC	TGTTGTTTGC	CCCTCCCCCG	TGCCTTCCTT	GACCTGGAA	GGTGCCACTC	3300
CCACTGTCCT	TTCCTAATAA	AATGAGGAAA	TTGCATCGCA	TTGTCTGAGT	AGGTGTCATT	3360
CTATTCTGGG	GGGTGGGGTG	GGGCAGGACA	GCAAGGGGGA	GGATTGGAA	GACAATAGCA	3420
GGCATGCTGG	GGATGCGGTG	GGCTCTATGG	AACCAGCTGG	GGCTCGACAG	CGCTGGATCT	3480
CCCGATCCCC	AGCTTTGCTT	CTCAATTCT	TATTCGATA	ATGAGAAAAA	AAGGAAAATT	3540
AATTTTAACA	CCAATTCAAGT	AGTTGATTGA	GCAAATGCGT	TGCCAAAAG	GATGCTTTAG	3600

□ □ □ □ □ □ □ □

CCCATTGCAC CACCAAGCGA AACATCGCAT CGAGCGAGCA CGTACTCGGA TGGAGGCCGG	5700
TCTTGTGAT CAGGATGATC TGGACGAAGA GCATCAGGGG CTCGCGCCAG CCGAACTGTT	5760
CGCCAGGCTC AAGGCGCGCA TGCCCGACGG CGAGGATCTC GTCGTGACCC ATGGCGATGC	5820
CTGCTTGCCG AATATCATGG TGGAAAATGG CCGCTTTCT GGATTCATCG ACTGTGGCCG	5880
GCTGGGTGTG GCGGACCGCT ATCAGGACAT AGCGTTGGCT ACCCGTGATA TTGCTGAAGA	5940
GCTTGGCGGC GAATGGGCTG ACCGCTTCCT CGTGCTTTAC GGTATGCCG CTTCCGATT	6000
CGCAGCGCAT CGCCTTCTAT CGCCTTCTT ACAGAGTTCTT CTGAGCGGGGA CTCTGGGTT	6060
CGAAATGACC GACCAAGCGA CGCCCAACCT GCCATCACGA GATTTCGATT CCACCGCCGC	6120
CTTCTATGAA AGGTTGGGCT TCGGAATCGT TTTCCGGGAC GCCGGCTGGA TGATCCTCCA	6180
GCGCGGGGAT CTCATGCTGG AGTTCTTCGC CCACCCCAAC TTGTTTATTG CAGCTTATAA	6240
TGGTTACAAA TAAAGCAATA GCATCACAAA TTTCACAAAT AAAGCATTAA TTTCACTGCA	6300
TTCTAGTTGT GGTTTGTCCA AACTCATCAA TCTATCTTAT CATGTCTGGA TCGCGGCCGC	6360
GATCCCCTCG AGAGCTTGGC GTAATCATGG TCATAGCTGT TTCTGTGTG AAATTGTTAT	6420
CCGCTCACAA TTCCACACAA CATAcgAGCC GGAGCATAAA GTGAAAGCC TGGGGTGCCT	6480
AATGAGTGAG CTAACTCACA TTAATTGCGT TCGCTCACT GCCCGCTTTC CAGTCGGAA	6540
ACCTGTCTG CCAGCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC GGTTTGCCTA	6600
TTGGGCGCTC TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTGTG CCGCTGCCG	6660
GAGCGGTATC AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAACG	6720
CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAA AAGGCCGCGT	6780
TGCTGGCGTT TTTCCATAGG CTCCGCCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA	6840
GTCAGAGGTG GCGAAACCCG ACAGGACTAT AAAGATACCA GGCCTTCCC CCTGAAAGCT	6900
CCCTCGTGCCT CTCTCCTGTT CCGACCCCTGC CGCTTACCGG ATACCTGTCC GCCTTCTCC	6960
CTTCGGGAAG CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG	7020
TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCGAC CGCTGCCCT	7080
TATCCGGTAA CTATCGTCTT GAGTCCAACC CGGTAAAGACA CGACTTATCG CCACTGGCAG	7140
CAGCCACTGG TAACAGGATT AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA	7200
AGTGGTGGCC TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA	7260
AGCCAGTTAC CTTGGAAAAA AGAGTTGGTA GCTCTTGATC CGGAAACAA ACCACCGCTG	7320
GTAGCGGTGG TTTTTTGTG TGCAAGCAGC AGATTACCGG CAGAAAAAAA GGATCTCAAG	7380
AAGATCCTTT GATCTTTCT ACGGGGCTG ACGCTCAGTG GAACGAAAC TCACGTTAAG	7440
GGATTTGGT CATGAGATTA TCAAAAGGA TCTTCACCTA GATCCTTTA AATTAAAAAT	7500
GAAGTTTAA ATCAATCTAA AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT	7560
TAATCACTGA GGCACCTATC TCAGCGATCT GTCTATTGCG TTCATCCATA GTTGCCTGAC	7620
TCCCCGTGCT GTAGATAACT ACGATAACGGG AGGGCTTACC ATCTGGCCCC AGTGTGCAA	7680

TGATACCGCG AGACCCACGC TCACCGGCTC CAGATTATC AGCAATAAAC CAGCCAGCCG	7740
GAAGGGCCGA GCGCAGAAGT GGTCCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAATT	7800
GTTGCCGGGA AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGAAC GTTGTTGCCA	7860
TTGCTACAGG CATCGTGGTG TCACGCTCGT CGTTTGGTAT GGCTTCATTG AGCTCCGGTT	7920
CCCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG CAAAAAAAGCG GTTAGCTCCT	7980
TCGGTCCTCC GATCGTTGTC AGAAGTAAGT TGGCCGCAGT GTTATCACTC ATGGTTATGG	8040
CAGCACTGCA TAATTCTCTT ACTGTCAATGC CATCCGTAAG ATGCTTTCT GTGACTGGTG	8100
AGTACTCAAC CAAGTCATTG TGAGAATAGT GTATGCCGCG ACCGAGTTGC TCTTGCCCCGG	8160
CGTCAATAACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC ATCATTGGAA	8220
AACGTTCTTC GGGCGAAAAA CTCTCAAGGA TCTTACCGCT GTTGAGATCC AGTTCGATGT	8280
AACCCACTCG TGCACCCAAC TGATCTTCAG CATCTTTAC TTTCACCCAGC GTTTCTGGGT	8340
GAGCAAAAAC AGGAAGGCAA AATGCCGAA AAAAGGGAAT AAGGGCGACA CGGAAATGTT	8400
GAATACTCAT ACTCTTCCTT TTTCAATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA	8460
TGAGCGGATA CATATTGAA TGTATTAGA AAAATAAACAA AATAGGGTT CCGCGCACAT	8520
TTCCCCGAAA AGTGCCACCT	8540

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9209 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GACGTCCGGG CCGCTCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG	60
AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGCATGGGGC	120
GGAGAATGGG CGGAACCTGGG CGGAGTTAGG GGCAGGGATGG GCGGAGTTAG GGGCGGGACT	180
ATGGTTGCTG ACTAATTGAG ATGCATGCTT TGCATACTTC TGCCTGCTGG GGAGCCTGGG	240
GACTTTCCAC ACCTGGTTGC TGACTAATTG AGATGCATGC TTTGCATACT TCTGCCTGCT	300
GGGGAGCCTG GGGACTTTCC ACACCTAAC TGACACACAT TCCACAGAAC TAATTCCCCT	360
AGTTATTAAT AGTAATCAAT TACGGGTCA TTAGTCATA GCCCATATAT GGAGTTCCGC	420
GTTACATAAC TTACGGTAAA TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG	480
ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA	540
TGGGTGGACT ATTTACGGTA AACTGCCAC TTGGCAGTAC ATCAAGTGTAA TCATATGCCA	600

AGTACGCCCT	CTATTGACGT	CAATGACGGT	AAATGGCCCG	CCTGGCATT	TGCCCAAGTAC	660
ATGACCTTAT	GGGACTTTCC	TACTTGGCAG	TACATCTACG	TATTAGTCAT	CGCTATTACC	720
ATGGTGATGC	GGTTTGGCA	GTACATCAAT	GGGCGTGGAT	ACCGGTTGA	CTCACGCGGA	780
TTTCCAAGTC	TCCACCCCCAT	TGACGTCAAT	GGGAGTTTGT	TTTGGCACCA	AAATCAACGG	840
GACTTTCAA	AATGTCGTA	AACTCCGCC	CCATTGACGC	AAATGGGCGG	TAGGCGTGT	900
CGGTGGGAGG	TCTATATAAG	CAGAGCTGGG	TACGTGAACC	GTCAGATCGC	CTGGAGACGC	960
CATCACAGAT	CTCTCACTAT	GGATTTTCAG	GTGCAGATT	TCAGCTTCCT	GCTAATCAGT	1020
GCTTCAGTCA	TAATGTCCAG	AGGACAAATT	GTTCTCTCCC	AGTCTCCAGC	AATCCTGTCT	1080
GCATCTCCAG	GGGAGAAGGT	CACAATGACT	TGCAGGGCCA	GCTCAAGTGT	AAGTTACATC	1140
CACTGGTCC	AGCAGAACGCC	AGGATCCTCC	CCCAAACCT	GGATTTATGC	CACATCCAAC	1200
CTGGCTTCTG	GAGTCCTCTG	TCGCTTCAGT	GGCAGTGGGT	CTGGGACTTC	TTACTCTCTC	1260
ACAATCAGCA	GAGTGGAGGC	TGAAGATGCT	GCCACTTATT	ACTGCCAGCA	GTGGACTAGT	1320
AAACCCACCCA	CGTTGGAGG	GGGGACCAAG	CTGGAAATCA	AACGTACGGT	GGCTGCACCA	1380
TCTGTCTCA	TCTTCCGCC	ATCTGATGAG	CAGTTGAAAT	CTGGAAC	CTCTGTTGT	1440
TGCCTGCTGA	ATAACTTCTA	TCCCAGAGAG	GCCAAAGTAC	AGTGGAAAGGT	GGATAACGCC	1500
CTCCAATCGG	GTAAC	CCCCCA	GGAGAGTGT	ACAGAGCAGG	ACAGCAAGGA	1560
AGCCTCAGCA	GCACCC	TGAC	GCTGAGCAA	GCAGACTACG	AGAAACACAA	1620
TGCGAAGTCA	CCC	CATCAGGG	CCTGAGCTCG	CCC	GTCACCAA	1680
TGTTGAATT	AGATCCGTTA	ACGGTTACCA	ACTACCTAGA	CTGGATT	CGTACATGC	1740
GGCCGTGATA	TCTACGTATG	ATCAGCCTCG	ACTGTGCCTT	CTAGTTGCCA	GCCATCTGTT	1800
GT	TTGCCCCT	CCCCCGTGCC	TTCC	CTGAAAGGT	CCACTCCCAC	1860
TAATAAAATG	AGGAAATTGC	ATCGCATTGT	CTGAGTAGGT	GTCATTCTAT	TCTGGGGGGT	1920
GGGGTGGG	AGGACAGCAA	GGGGGAGGAT	TGGAAGACA	ATAGCAGGCA	TGCTGGGGAT	1980
GCGGTGGG	CTATGGAACC	AGCTGGG	CGACAGCTAT	GCCAA	GTACAGTACG	2040
ACGTCAATGA	CGGTAAATGG	CCC	GCTGGC	ATTATGCCA	GTACATGACC	2100
TTCCTACTTG	GCAGTACATC	TACGTATTAG	TCATCGCTAT	TACCATGGT	ATGCGGTTT	2160
GGCAGTACAT	CAATGGCGT	GGATAGCGT	TTGACTCAGG	GGGATT	TCCAAGTCCACC	2220
CCATTGACGT	CAATGGGAGT	TTGTTTGGC	ACCAAAATCA	ACGGGACTTT	CCAAAATGTC	2280
GTAACAACTC	CGCCCCATTG	ACGCAAATGG	GCGGTAGGG	TGTACGGTGG	GAGGTCTATA	2340
TAAGCAGAGC	TGGGTACGTC	CTCACATTCA	GTGATCAGCA	CTGAACACAG	ACCCGTGAC	2400
ATGGGTTGGA	GCCTCATCTT	GCTCTTCC	GTGCGTGT	CTACCGGTGT	CCTGTCCCAG	2460
GTACAAC	GTG	GGCTGCTGG	GGCTGAGCTG	GTGAAGCCTG	GGGCCTCAGT	2520
TGCAAGG	CTGGCTACAC	ATTTACCA	GACTACATGC	ACTGGTAAA	ACAGACACCT	2580
GGTCGGG	GGCAATGGAT	TGGAGCTATT	TATCCCGGAA	ATGGTGATAC	TTCCTACAAT	2640

200 150 100 50 0

CAGAAGTTCA AAGGCAAGGC	ATTGACT GCAGACAAAT CCTCCAGCAC	CTACATG	2700
CAGCTCAGCA GCCTGACATC TGAGGACTCT GCGGTCTATT ACTGTGCAAG ATCGACTTAC			2760
TACGGCGGTG ACTGGTACTT CAATGTCTGG GGCGCAGGGA CCACGGTCAC CGTCTCTGCA			2820
GCTAGCACCA AGGGCCCATC GGTCTCCCC CTGGCACCC CTCACAAGAG CACCTCTGGG			2880
GGCACAGCGG CCCTGGGCTG CCTGGTCAAG GACTACTCC CCGAACCGGT GACGGTGTG			2940
TGGAACTCAG GCGCCCTGAC CAGCGCGTG CACACCTCC CGGCTGTCT ACAGTCCTCA			3000
GGACTCTACT CCCTCAGCAG CGTGGTGACC GTGCCCTCCA GCAGCTTGGG CACCCAGACC			3060
TACATCTGCA ACGTGAATCA CAAGCCCAGC AACACCAAGG TGGACAAGAA AGCAGAGCCC			3120
AAATCTTGTG ACAAAAATCA CACATGCCA CCGTGCCAG CACCTGAAC CCTGGGGGGA			3180
CCGTCAGTCT TCCTCTTCCC CCCAAAACCC AAGGACACCC TCATGATCTC CGGGACCCCT			3240
GAGGTACAT GCGTGGTGGT GGACGTGAGC CACGAAGACC CTGAGGTCAA GTTCAACTGG			3300
TACGTGGACG GCGTGGAGGT GCATAATGCC AAGACAAAGC CGCGGGAGGA GCAGTACAAC			3360
AGCACGTACC GTGTGGTCAG CGTCCTCACC GTCTGCACC AGGACTGGCT GAATGGCAAG			3420
GAGTACAAGT GCAAGGTCTC CAACAAAGCC CTCCAGCCC CCATCGAGAA AACCATCTCC			3480
AAAGCCAAAG GGCAGCCCCG AGAACACACAG GTGTACACCC TGCCCCCATC CGGGATGAG			3540
CTGACCAAGA ACCAGGTCAAG CCTGACCTGC CTGGTCAAAG GCTTCTATCC CAGCGACATC			3600
GCCGTGGAGT GGGAGAGCAA TGGGCAGCCG GAGAACAACT ACAAGACAC GCCTCCCGTG			3660
CTGGACTCCG ACGGCTCCTT CTTCCCTAC AGCAAGCTCA CGTGGACAA GAGCAGGTGG			3720
CAGCAGGGGA ACGTCTTCTC ATGCTCCGTG ATGCATGAGG CTCTGCACAA CCACTACACG			3780
CAGAAGAGCC TCTCCCTGTC TCCGGTAAA TGAGGATCCG TTAACGGTTA CCAACTACCT			3840
AGACTGGATT CGTGACAACA TGCAGCCGTG ATATCTACGT ATGATCAGCC TCGACTGTGC			3900
CTTCTAGTTG CCAGCCATCT GTTGTGTTGCC CCTCCCCGT GCCTTCCTTG ACCCTGGAAG			3960
GTGCCACTCC CACTGTCTT TCCTAATAAA ATGAGGAAAT TGCATCGCAT TGTCTGAGTA			4020
GGTGTCAATT TATTCTGGGG GGTGGGTGG GGCAGGACAG CAAGGGGAG GATTGGGAAG			4080
ACAATAGCAG GCATGCTGGG GATGCGGTGG GCTCTATGGA ACCAGCTGGG GCTCGACAGC			4140
GCTGGATCTC CCGATCCCCA GCTTGCCTTC TCAATTCTT ATTGCTAA TGAGAAAAAA			4200
AGGAAAATTA ATTTAACAC CAATTCAAGTA GTTGATTGAG CAAATGGTT GCCAAAAAGG			4260
ATGCTTTAGA GACAGTGTTC TCTGCACAGA TAAGGACAAA CATTATTCAAG AGGGAGTACC			4320
CAGAGCTGAG ACTCCTAACGC CAGTGAGTGG CACAGCATTG TAGGGAGAAA TATGCTTGTC			4380
ATCACCGAAG CCTGATTCCG TAGAGCCACA CCTGGTAAAG GGCCAACTG CTCACACAGG			4440
ATAGAGAGGG CAGGAGCCAG GGCAGAGCAT ATAAGGTGAG GTAGGATCAG TTGCTCCTCA			4500
CATTTGCTTC TGACATAGTT GTGTTGGGAG CTTGGATAGC TTGGACAGCT CAGGGCTGCG			4560
ATTTCGCGCC AAACTTGACG GCAATCCTAG CGTGAAGGCT GGTAGGATT TATCCCCGCT			4620
GCCATCATGG TTGACCATG GAACTGCATC GTGCCGTGT CCCAAAATAT GGGGATTGGC			4680

AAGAACGGAG ACCTACCCCTG	CTCCGCTC AGGAACGAGT TCAAGTACTT	AAAGAATG	4740
ACCACAAACCT CTTCACTGGG	AGGTAAACAG AATCTGGTGA TTATGGGTAG	GAAAACCTGG	4800
TTCTCCATTG CTGAGAAGAA	TCGACCTTA AAGGACAGAA	TAAATATAAGT TCTCAGTAGA	4860
GAACCTCAAAG AACCAACCACG	AGGAGCTCAT TTTCTGCCA	AAAGTTGGTA TGATGCCCTA	4920
AGACTTATTG AACAAACCGGA	ATTGGCAAGT AAAGTAGACA	TGGTTGGAT AGTCGGAGGC	4980
AGTTCTGTTT ACCAGGAAGC	CATGAATCAA CCAGGCCACC	TTAGACTCTT TGTGACAAGG	5040
ATCATGCAGG AATTGAAAG	TGACACGTTT TTCCCAGAAA	TTGATTGGGG GAAATATAAA	5100
CTTCTCCAG AATAACCCAGG	CGTCCTCTCT GAGGTCCAGG	AGGAAAAAGG CATCAAGTAT	5160
AAGTTTGAAG TCTACGAGAA	GAAAGACTAA CAGGAAGATG	CTTCAAGTT CTCTGCTCCC	5220
CTCCTAAAGC TATGCATTTT	TATAAGACCA TGGGACTTTT	GCTGGCTTA GATCAGCCTC	5280
GACTGTGCCT TCTAGTTGCC	AGCCATCTGT TGTGCCCC	TCCCCCGTGC CTTCCTTGAC	5340
CCTGGAAAGGT GCCACTCCCA	CTGTCCTTTC CTAATAAAAT	GAGGAAATTG CATCGCATTG	5400
TCTGAGTAGG TGTCAATTCTA	TTCTGGGGGG TGGGGTGGGG	CAGGACAGCA AGGGGGAGGA	5460
TTGGGAAGAC AATAGCAGGC	ATGCTGGGA TGCGGTGGC	TCTATGGAAC CAGCTGGGGC	5520
TCGAGCTACT AGCTTTGCTT	CTCAATTCT TATTTGCATA	ATGAGAAAAA AAGGAAAATT	5580
AATTTAACCA CCAATTCACT	AGTTGATTGA GCAAATGCGT	TGCCAAAAG GATGCTTTAG	5640
AGACAGTGTG CTCTGCACAG	ATAAGGACAA ACATTATTCA	GAGGGAGTAC CCAGAGCTGA	5700
GACTCCTAAG CCAGTGAGTG	GCACAGCATT CTAGGGAGAA	ATATGCTTGT CATCACCGAA	5760
GCCTGATTCC GTAGAGCCAC	ACCTTGGTAA GGGCCAATCT	GTCACACAG GATAGAGAGG	5820
GCAGGAGCCA GGGCAGAGCA	TATAAGGTGA GGTAGGATCA	GTTGCTCCTC ACATTGCTT	5880
CTGACATAGT TGTGTTGGGA	GCTTGGATCG ATCCTCTATG	GTTGAACAAG ATGGATTGCA	5940
CGCAGGTTCT CCGGCCGCTT	GGGTGGAGAG GCTATTGGC	TATGACTGGG CACAACAGAC	6000
AATCGGCTGC TCTGATGCCG	CCGTGTTCCG GCTGTCAGCG	CAGGGCGGCC CGGTTCTTT	6060
TGTCAAGACC GACCTGTCCG	GTGCCCTGAA TGAAC TGCACTGCAG	GACGAGGCAG CGCGGCTATC	6120
GTGGCTGGCC ACGACGGCG	TTCCTGCGC AGCTGTGCTC	GACGTTGTCA CTGAAGCGGG	6180
AAGGGACTGG CTGCTATTGG	GCGAAGTGCC GGGCAGGAT	CTCCTGTCACT CTCACCTTGC	6240
TCCTGCCGAG AAAGTATCCA	TCATGGCTGA TGCAATGCCG	CGGCTGCATA CGCTTGATCC	6300
GGCTACCTGC CCATTGACC	ACCAAGCGAA ACATCGCATC	GAGCGAGCAC GTACTCGGAT	6360
GGAAGCCGGT CTTGTCGATC	AGGATGATCT GGACGAAGAG	CATCAGGGGC TCGCGCCAGC	6420
CGAAACTGTTC GCCAGGCTCA	AGGCCGCAT GCGGACGGC	GAGGATCTCG TCGTGACCCA	6480
TGGCGATGCC	TGCTTGCCGA ATATCATGGT	GGAAAATGGC CGCTTTCTG GATTCACTGA	6540
CTGTGGCCGG CTGGGTGTGG	CGGACCGCTA TCAGGACATA	GGCTGGCTA CCCGTGATAT	6600
TGCTGAAGAG CTTGGCGGG	AATGGGCTGA CCGCTTCCTC	GTGCTTTACG GTATCGCCGC	6660
TCCCGATTG CAGCGCATCG	CCTTCTATCG CCTTCTTGAC	GAGTTCTTCT GAGCGGGACT	6720

CTGGGGTTCG AAATGACCGA	AGCGACG CCCAACCTGC CATCACGAGA	CGATTCC	6780
ACCGCCGCCT TCTATGAAAG GTTGGGCTTC GGAATCGTTT TCCGGGACGC CGGCTGGATG			6840
ATCCTCCAGC GCAGGGATCT CATGCTGGAG TTCTTCGCC ACCCCAACCT GTTTATTGCA			6900
GCTTATAATG GTTACAAATA AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT			6960
TCACTGCATT CTAGTTGTGG TTTGTCCAAA CTCATCATC TATCTTATCA TGCTCTGGATC			7020
GCAGCCGGGA TCCCCTCGAG AGCTTGGCGT AATCATGGTC ATAGCTGTTT CCTGTGTGAA			7080
ATTGTTATCC GCTCACAAATT CCACACAAACA TAGGAGCCGG AACGCATAAAG TGTAAGGCCT			7140
GGGGTGCCTA ATGAGTGAGC TAACTCACAT TAATTGCGTT GCGCTCACTG CCCGCTTCC			7200
AGTCGGGAAA CCTGTCGTGC CAGCTGCATT AATGAATCGG CCAACGGCGC GGGAGAGGCG			7260
GTTTGCAT TGGCGCTCT TCCGCTTCCT CGCTCACTGA CTCGCTGCCTC TCGGTGTTTC			7320
GGCTGCAGCG AGCGGTATCA GCTCACTCAA AGGGCGTAAT ACGGTTATCC ACAGAAATCAG			7380
GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA			7440
AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCGG TGACGAGCAT CACAAAAATC			7500
GACGCTCAAG TCAGAGGTGG CGAAACCGA CAGGACTATA AAGATACCAG GCGTTTCCCC			7560
CTGGAAGCTC CCTCGTGCCTC TCTCCTGTT CGACCCCTGCC GCTTACCGGA TACCTGTCCG			7620
CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCATGCTC ACGCTGTAGG TATCTCAGTT			7680
CGGTGTAGGT CGTTGCTCC AAGCTGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC			7740
GCTGCGCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC			7800
CACTGGCAGC AGCCACTGGT AACAGGAGTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG			7860
AGTTCTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAC GACAGTATTT GGTATCTGCG			7920
CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA			7980
CCACCGCTGG TAGCGGTGGT TTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG			8040
GATCTCAAGA AGATCCTTTG ATCTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAAC			8100
CACGTTAAGG GATTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTAA			8160
ATTAAAAATG AAGTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT			8220
ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTCGT TCATCCATAG			8280
TTGCCTGACT CCCCGTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA			8340
GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTATCA GCAATAAAC			8400
AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT			8460
CTATTAATTG TTGCCGGAA GCTAGAGTAA GTAGTTGCC AGTTAATAGT TTGCGCAACG			8520
TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA			8580
GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAGCGG			8640
TTAGCTCCTT CGGTCCCTCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA			8700
TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTCTG			8760

TGACTGGTGA GTACTCAACC	TCATTCT GAGAATAGTG TATGCGGCGA	AGTTGCT	8820
CTTGCCTGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA			8880
TCATTGGAAA ACGTTCTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA			8940
GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTACT TTCACCAGCG			9000
TTTCTGGGTG AGCAAAAACA GGAAGGCAAATGCCGAAA AAAGGGATA AGGGCGACAC			9060
GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT			9120
ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTTC			9180
CGCGCACATT TCCCCGAAAAA GTGCCACCT			9209

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 47 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATCACAGATC TCTCACCATG GATTTTCAGG TGCAGATTAT CAGCTTC

47

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: YES
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TGCAGCATCC GTACGTTGA TTTCCAGCTT

30

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 384 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..384

(ix) FEATURE:

(A) NAME/KEY: mat\_peptide

(B) LOCATION: 67..384

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATG GAT TTT CAG GTG CAG ATT ATC AGC TTC CTG CTA ATC AGT GCT TCA Met Asp Phe Gln Val Gln Ile Ile Ser Phe Leu Leu Ser Ala Ser -22 -20 -15 -10	48
GTC ATA ATG TCC AGA GGG CAA ATT GTT CTC TCC CAG TCT CCA GCA ATC Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile -5 -1 1 5 10	96
CTG TCT GCA TCT CCA GGG GAG AAG GTC ACA ATG ACT TGC AGG GCC AGC Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser 15 20 25	144
TCA AGT GTA AGT TAC ATC CAC TGG TTC CAG CAG AAG CCA GGA TCC TCC Ser Ser Val Ser Tyr Ile His Trp Phe Gln Gln Lys Pro Gly Ser Ser 30 35 40	192
CCC AAA CCC TGG ATT TAT GCC ACA TCC AAC CTG GCT TCT GGA GTC CCT Pro Lys Pro Trp Ile Tyr Ala Thr Ser Asn Leu Ala Ser Gly Val Pro 45 50 55	240
GTT CGC TTC AGT GGC AGT GGG TCT GGG ACT TCT TAC TCT CTC ACA ATC Val Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile 60 65 70	288
AGC AGA GTG GAG GCT GAA GAT GCT GCC ACT TAT TAC TGC CAG CAG TGG Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp 75 80 85 90	336
ACT AGT AAC CCA CCC ACG TTC GGA GGG GGG ACC AAG CTG GAA ATC AAA Thr Ser Asn Pro Pro Thr Phe Gly Gly Thr Lys Leu Glu Ile Lys 95 100 105	384

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCGGCTCCCA CGCGTGTCT GTCCCCAG

27

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: YES
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: 3  
(D) OTHER INFORMATION: /note= "Nucleotide 3 is N wherein N is G or C."
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: 18  
(D) OTHER INFORMATION: /note= "Nucleotide 18 is N wherein N is A or C."
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: 19  
(D) OTHER INFORMATION: /note= "Nucleotide 19 is N wherein N is A or G."
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: 25  
(D) OTHER INFORMATION: /note= "Nucleotide 25 is N wherein N is G or A."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGNTGTTGTG CTAGCTGNNG AGACNGTGA

29

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 420 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 1..420
- (ix) FEATURE:  
(A) NAME/KEY: mat\_peptide  
(B) LOCATION: 58..420

QW  
BS

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATG GGT TGG AGC CTC ATC TTG CTC TTC CTT GTC GCT GTT GCT ACG CGT Met Gly Trp Ser Leu Ile Leu Leu Phe Leu Val Ala Val Ala Thr Arg	48
-19 -15 -10 -5	
GTC CTG TCC CAG GTA CAA CTG CAG CAG CCT GGG GCT GAG CTG GTG AAG Val Leu Ser Gln Val Gln Leu Gln Pro Gly Ala Glu Leu Val Lys	96
-1 1 5 10	
CCT GGG GCC TCA GTG AAG ATG TCC TGC AAG GCT TCT GGC TAC ACA TTT Pro Gly Ala Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe	144
15 20 25	
ACC AGT TAC AAT ATG CAC TGG GTA AAA CAG ACA CCT GGT CGG GGC CTG Thr Ser Tyr Asn Met His Trp Val Lys Gln Thr Pro Gly Arg Gly Leu	192
30 35 40 45	
GAA TGG ATT GGA GCT ATT TAT CCC GGA AAT GGT GAT ACT TCC TAC AAT Glu Trp Ile Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn	240
50 55 60	
CAG AAG TTC AAA GGC AAG GCC ACA TTG ACT GCA GAC AAA TCC TCC AGC Gln Lys Phe Lys Gly Lys Ala Thr Leu Thr Ala Asp Lys Ser Ser Ser	288
65 70 75	
ACA GCC TAC ATG CAG CTC AGC AGC CTG ACA TCT GAG GAC TCT GCG GTC Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val	336
80 85 90	
TAT TAC TGT GCA AGA TCG ACT TAC TAC GGC GGT GAC TGG TAC TTC AAT Tyr Tyr Cys Ala Arg Ser Thr Tyr Tyr Gly Gly Asp Trp Tyr Phe Asn	384
95 100 105	
GTC TGG GGC GCA GGG ACC ACG GTC ACC GTC TCT GCA Val Trp Gly Ala Gly Thr Thr Val Thr Val Ser Ala	420
110 115 120	